

**Network Traffic Modelling and Router
Performance Optimization Using
Fuzzy Logic and Genetic Algorithms**

BY

Abul Kashem Muhammed Ziaur Rahman.

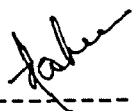
THESIS

Submitted to the Faculty of IT
University of Technology Sydney (UTS)
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Computing Sciences

"I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis."



Zia A. Rahman
Sydney, Australia.

Abstract

Accurate computer network traffic models are required for many network tasks such as network analysis, performance optimization and areas of traffic engineering such as avoiding congestion or guaranteeing a specific quality of service (QoS) to an application. Existing traffic modelling techniques rely on precise mathematical analysis of extensive measured data such as packet arrival time, packet size and server-side or client-side round trip time. With the advent of high speed broadband networks, gathering an acceptable quantity of data needed for the precise representation of traffic is a difficult, time consuming, expensive and in some cases almost an impossible task. A possible alternative is to employ fuzzy logic based models which can represent processes characterized by imprecise data, which is generally easier to gather. The effectiveness of these models has been demonstrated in many industrial applications. This work develops fuzzy logic based traffic models using imprecise data sets that can be obtained realistically. Optimizing the performance of a router requires the optimization of a number of conflicting objectives. A possible approach is to express it as a multi-objective problem. Multi-objective evolutionary algorithms (MOEA) can be used for solving such problems. This research proposes two fuzzy logic based traffic models: fuzzy group model and fuzzy state model. These models together with MOEA are used to propose a simple and fast router buffer management scheme. The developed fuzzy group model includes a parameter which is also useful for measuring the irregular traffic patterns known as burstiness. The experimental results are promising.

Acknowledgments

This thesis is the outcome of years of hard work. During this long journey, I have been supported and motivated by many people. It is a pleasant feeling that now I have the opportunity to express my gratitude for all of them.

Foremost, I would like to express my deep gratitude to my principal supervisor, Professor Jenny Edwards. She has shown me the way of research. She has always been available to advise me. Besides of being an excellent supervisor, she is a great motivator. During the past five years, Jenny was as close as a relative or good friend to me. Without her constant support this thesis would never come to light.

I would like to thank my co-supervisor Dr Paul Kennedy who kept a constant eye on my progress. He helped me immensely to develop my writing and communication skills. He was always my last resort in solving the most difficult problems. He could not even realize how much I have learned from him.

Special thanks to Dr Andrew Simmonds, my other co-supervisor. His immense knowledge on computer networks and cordial but valuable comments made me look inside of the complex area of computer networks.

I feel a profound sense of gratitude for my late father (whom I lost during the middle of my research) and mother who taught me all the good things that really matter in life. The happy memory of my father is still the persistent source of inspiration for my journey in this life. I am grateful to my brother and sisters. I am glad to be one of them.

I am very grateful to my wife Curie, for her love and patience during the last five years. One of the greatest moments that we had in this period was the birth of our only child, Ryan, who provided a new joyful dimension to our life. Thanks Ryan, life would be incomplete without you.

*Dedicated to Ryan M Rahman
My only son (and sun)*

Contents

1	Introduction	1
2	Historical View of Network Traffic Models	5
2.1	Internet Traffic	5
2.2	Application of Traffic Models	6
2.3	Mathematical Representation of Traffic Streams	8
2.4	First Generation Models - Non-Self-Similar Models	9
2.4.1	Poisson Models of Internet traffic	9
2.4.2	Markovian Models	10
2.5	The Discovery of Self-Similarity	11
2.5.1	Stationarity	12
2.5.2	Self-Similarity	13
2.5.2.1	Hurst parameter estimation	14
2.5.2.2	Generating Self-Similar Traffic	15
2.5.3	Evidence for Self-Similarity in Network Traffic	16
2.5.4	Failure of Poisson Based First Generation Models	18
2.6	Second Generation Models - Self-Similar Models	19
2.6.1	Chaotic Map Model	19
2.6.2	Brownian Motion Model	21
2.6.3	Discussion of Self-similar Models	22
2.7	Evidence of Non-Stationarity	22
2.8	Third Generation Models	23
2.8.1	The Fractional Sum Difference-Moving Average (FSD-MA) Model	24
2.9	The Need for a New Approach	25
2.10	Measuring Burstiness of Network Traffic	27
2.10.1	Energy Plots	28
2.11	Conclusion	30

3	Network Performance and Management	31
3.1	Network Performance	31
3.1.1	Reasons for Performance Degradation	32
3.1.2	Performance Objectives	33
3.2	Buffer Dimensioning	34
3.3	Queue Management and Fairness	36
3.3.1	Passive Queue Management	37
3.3.2	Active Queue Management	37
3.3.2.1	Random Early Detection	37
3.3.2.2	Stabilized Random Early Detection	39
3.3.2.3	Fair Random Early Detection	40
3.4	Network Traffic Measurement	41
3.4.1	Content Non-aware Measurement	41
3.4.2	Content-aware Measurement	42
3.4.3	Active vs Passive Measurement	42
3.5	Router Performance Optimization: A Multi-Objective Optimiza- tion Problem	43
3.5.1	Multi-Objective Optimization Problems	43
3.5.2	Router Performance Optimization as an MO Problem . . .	44
3.6	Design Principles for a Queue Management Scheme	45
3.7	Conclusion	47
4	Soft Computing	48
4.1	Fuzzy Logic	48
4.1.1	Fuzzy Sets, Membership Functions and Linguistic Variables	49
4.1.2	Fuzzy Logic Systems	51
4.1.2.1	Fuzzification	51
4.1.2.2	Fuzzy Inference	53
4.1.2.3	Defuzzification	54
4.2	Approaches to Solving MO Problems	56
4.2.1	Pareto Optimality	57
4.2.2	Classical Approach	59
4.2.3	Evolutionary Approach	60
4.2.3.1	MOEA: Aggregating Approach	61
4.3	Review of the Router Optimization Problem	62
4.4	Conclusion	63

5	Fuzzy Logic Based Traffic Models	64
5.1	Introduction	64
5.2	The Fuzzy State Model	64
5.2.1	Construction of Difference Windows	67
5.2.2	Fuzzification of Difference Windows and Packet Counts . .	68
5.2.3	Determination of Trend Fuzzy Sets	70
5.2.4	Determination of State Fuzzy Sets	73
5.2.5	Discussion of the State Model	77
5.3	The Fuzzy Group Model	77
5.3.1	Modelling a Flow	80
5.3.2	Modelling Aggregate Traffic	81
5.3.3	Determination of R	83
5.3.4	Discussion of the Group Model	87
5.4	Conclusion	88
6	Traffic Analysis Using Our Proposed Traffic Model	89
6.1	Introduction	89
6.2	Experiment 1: H and R Parameters for Real Traces	91
6.2.1	Determination of the H parameter	91
6.2.2	Determination of the R Parameter	92
6.2.3	Experimental Results and Discussion	93
6.2.3.1	Visual Summary of H and R Parameters	93
6.2.3.2	H and R values at different time of day and day of week	95
6.2.3.3	Stationarity	95
6.2.3.4	H and R as a function of the number of connections	96
6.2.3.5	H and R values as a function of link utilization .	100
6.3	Experiment 2: Analysis of Trace-Driven Queuing	101
6.3.1	Generating Self-Similar Traffic	101
6.3.2	The Trace-Driven Queue Simulator	102
6.3.3	Experimental Results and Discussion	103
6.3.3.1	H as an indicator of queuing performance	104
6.3.3.2	R as an indicator of queuing performance	106
6.4	Experiment 3: Analysis of Burstiness using Energy Plots	108
6.4.1	Goal of the Experiment	108
6.4.2	Experimental Data	108
6.4.3	Experimental Methods	109

6.4.4	Results	110
6.4.5	H as a measure of Burstiness	110
6.4.6	R as a measure of Burstiness	110
6.5	Conclusion	110
7	Router Performance Optimization Using Our Proposed Traffic Model	111
7.1	FAFRED (Fuzzy Adaptive Fair Random Early Detection)	111
7.2	Discussion of FAFRED	115
7.2.1	Discussion on Setting up Parameters for FAFRED	115
7.3	Simulation of FAFRED and Other AQM Schemes	117
7.3.1	Topology Used	118
7.3.2	Simulation Setup	118
7.3.3	Initial Parameters	118
7.3.4	Setting up the GA for FAFRED Optimization	119
7.3.5	Analysis of the Results	121
7.3.5.1	The throughput vs time plots	121
7.3.5.2	Queue length vs time	124
7.3.5.3	TCP Sequence Number vs Time	127
7.3.5.4	Fairness	129
7.4	Conclusion	129
8	Conclusions and Future Work	131
8.1	Conclusions	131
8.2	Future Research	133
	References	136
A	Glossary	149
B	List of Abbreviations	153
C	An example of a Genetic Algorithm	155
D	Graphs of Queuing Simulations	159
E	Pseudocodes and Scripts	162
E.1	Pseudocode to Compute R	162
E.2	Computing R Using Perl	163

E.2.1	The Script	163
E.2.2	Runtime	168
E.3	MATLAB script to compute the Hurst Parameter	168
E.4	Trace-Driven Queue Simulator	170
E.5	Perl Script to read 2005 AMP trace	171
E.6	Pseudocode for the Simulation of FAFRED	173
F	List of Symbols	175
G	Traces Used	177

List of Figures

2.1	User Traffic and Aggregate Traffic	6
2.2	Packet count vs time of self-similar network traffic (on left) and Poisson traffic (on right) at various time scales. (Reprinted with permission from authors of [79]). Self-similar traffic exhibits high variability at all time scales.	17
3.1	Drop Probability of Random Early Detection.	38
3.2	Drop Probability of Stabilized RED	39
4.1	Two Popular Shapes of Membership Functions. (a) Triangular (b) Trapezoidal	49
4.2	Fuzzy Sets Cold, Warm and Hot with Triangular Membership Function	50
4.3	The modules in a Fuzzy Logic based Control system. The Fuzzification module maps the crisp value to fuzzy sets. The Inference Engine produces fuzzy output based on a rulebase. The defuzzification module produces crisp output.	52
4.4	Fuzzy Inferencing and defuzzification of the Router Buffer Control Problem using min Inference and centroid defuzzification. The dashed line corresponds to the min operation, as in (a) $\min(0.4, 0.7)=0.4$ and in (b) $\min(0.6, 0.3)=0.3$. (c) is the defuzzification.	54
4.5	Inference rules for a Fuzzy Logic Controlled Router Buffer	55
4.6	The Pareto Optimal Set for the motor manufacturer problem. The cheapest solution makes the highest Pollution and the costliest solution makes the lowest pollution. The ideal solution (I) is unattainable.	58
5.1	Logical Block Diagram of the Fuzzy State Model	66
5.2	Construction of difference windows from the arriving packet counts.	67

5.3	Triangular Membership Function for POSITIVE, ZERO and NEG- ATIVE Fuzzy Sets for d_i and d'_i	68
5.4	Triangular Membership Function for SMALL, MEDIUM and LARGE fuzzy sets for N_i	69
5.5	Various Trend Fuzzy Sets	71
5.6	Membership functions and grade of membership for the fuzzifica- tion of {40,60,70} traffic stream	72
5.7	Various State Fuzzy States and Their Transitions	76
5.8	Packets, T-groups and f-groups.	79
5.9	Various Transformations in a Router. (a) Two T-groups arrive at a router. (b) A packet is dropped and other packets are queued (c) A new T-group is formed.	82
5.10	Fuzzy Logic System to Determine R	84
5.11	Membership Function of (a) B and (b) δ for aggregate traffic. . .	85
5.12	Triangular Membership Function of Output Fuzzy Sets for R: SMALL, MEDIUM and LARGE.	87
6.1	Hurst parameter of 2001, 2003 and 2003 traces at different time of the day and days of the week.	94
6.2	R Parameter of 2001, 2003 and 2005 traces at different time of day and days of week.	94
6.3	H Parameter vs Total No of Connections - 2001 Traces	96
6.4	H Parameter vs Total No of Connections - 2003 Traces	96
6.5	H Parameter vs Total No of Connections - 2005 Traces	97
6.6	R Parameter vs Total No of Connections - 2001 Traces	98
6.7	R Parameter vs Total No of Connections - 2003 Traces	98
6.8	R Parameter vs Total No of Connections - 2005 Traces	99
6.9	H Parameter as a Function of Link Utilization	100
6.10	R Parameter as a Function of Link Utilization	100
6.12	Queuing simulation of 2003 Auckland traffic	105
6.13	Queuing Simulation of the 2005 AMP Traffic	105
6.11	Queuing Simulation of the 2001 Auckland Traces	107
6.14	Queuing Simulation of Generated Self-similar Traffic.	107
6.15	Energy Plots of the Traces of Table 6.3	109

7.1	Block diagram of FAFRED. The CP module predicts congestion by monitoring traffic state and computes global drop function $r(t)$. The PD module does the actual dropping.	112
7.2	Topology Used in the Simulation of various AQM scheme.	117
7.3	Throughput (in bytes per second) vs Time (in second) Plots of (a) RED (b) FRED	122
7.4	Throughput (in bytes per second) vs Time (in second) Plots of (a) Drop-Tail and (b) FAFRED	123
7.5	Queue Length (in bytes) vs Time (in second) of (a) RED (b) FRED. The red lines are the average queue length and the blue lines are the instantaneous queue length.	125
7.6	(a) Queue Length (no of packets) vs Time (in second) for Drop-Tail. (b) Queue Length (in bytes) vs Time (in second) of FAFRED. The red line is the average queue length and the blue lines are the instantaneous queue length.	126
7.7	TCP Sequence Number vs Time (in seconds) Plots of (a) RED (b) FRED	127
7.8	TCP Sequence Number vs Time (in seconds) Plots of (a) Drop-Tail and (b) FAFRED	128
D.1	Queuing Performance of AMP05-26 Trace at Different Loads. . . .	159
D.2	Queuing Performance of AMP05-33 Trace at Different Loads. . . .	160
D.3	Queuing Performance of AMP05-15 Trace at Different Loads. . . .	160
D.4	Queuing Performance of AMP05-01 Trace at Different Loads. . . .	161

List of Tables

5.1	Rule-base for Fuzzy Inference Module to Determine the Trend Fuzzy Sets	70
5.2	Rule-base for Fuzzy Inferencing to Compute Output Fuzzy Sets for R	86
6.1	Traces Used for the Analysis.	91
6.2	Traces Used to Study Queuing	104
6.3	Traces Used in the Study of Burstiness.	108
7.1	Rule-base used by the CP module to compute the global drop probability.	113
7.2	Active Connections During FAFRED Simulation	118
7.3	Fairness Index of Various Queue Management Schemes.	129
E.1	Time Required to Computer R for off-line traffic tarces	168